DOCUMENT RESUME

ED 427 520 FL 025 692

AUTHOR Zampini, Mary L.
TITLE The Relationship between the Production and Perception of L2 Spanish Stops.
ISSN ISSN-0898-8471
PUB DATE 1998-00-00
NOTE 18p.; For the complete volume of working papers see FL 025 687.
PUB TYPE Journal Articles (080) -- Reports - Research (143)
JOURNAL CIT Texas Papers in Foreign Language Education; v3 n3 p85-100 Fall 1998
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Auditory Perception; Higher Education; Language Research; *Phonology; *Pronunciation; Second Language Learning; *Second Languages; *Spanish
IDENTIFIERS *English Speaking

ABSTRACT

A study explored the relationship between the second language (L2) production of the Spanish stop consonants /p/ and /b/. Subjects were native English-speaking students enrolled in an undergraduate Spanish phonetics course. Data were collected on their production of Spanish /p,b/ in sentence context, and on their perceptual boundary between the phonemes as reflected by changes in voice onset time (VOT). Mean production and perception VOTs were compared. Results do not reveal a strong correlation between learners' perceptual abilities and production of the L2 Spanish stops. In particular, while some learners' L2 productions approximated those of native Spanish speakers, their perceptual boundaries were similar to boundaries for monolingual English speakers, and vice versa. (MSE)

The Relationship Between the Production and Perception of L2 Spanish Stops

MARY L. ZAMPINI, The University of Arizona

The Relationship Between the Production and Perception of L2 Spanish Stops

MARY L. ZAMPINI, The University of Arizona

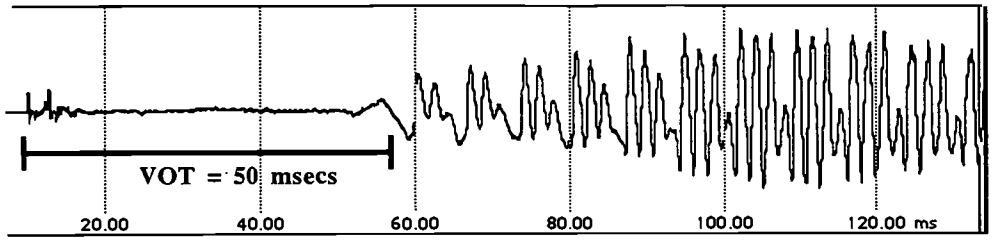
This paper explores the relationship between the second language (L2) production and perception of the Spanish stop consonants /p/ and /b/. An experiment was conducted that collected data on adult English-speaking learners' production of Spanish /p, b/ in a sentence context. The same learners also completed a series of perception experiments that examined their perceptual boundary between /p/ and /b/ as reflected by changes in voice onset time (VOT). The mean VOTs produced by the learners in the production experiment were compared to their perceptual VOT boundary. The results do not reveal a strong correlation between learners' perceptual capabilities and production of the L2 Spanish stops. In particular, while some learners' L2 productions approximated those of native Spanish speakers, their perceptual boundaries were similar to boundaries for monolingual English speakers, and vice versa.

INTRODUCTION

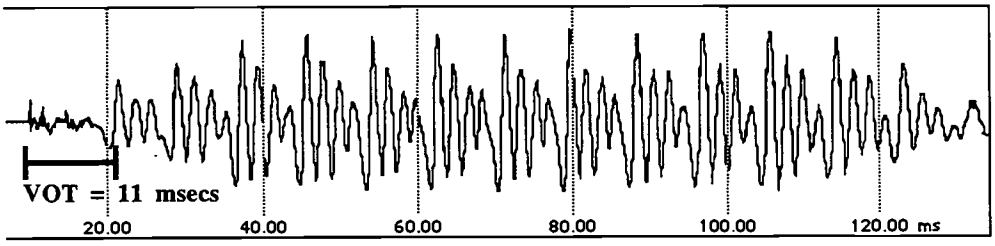
An important issue of second language (L2) pronunciation and phonological acquisition is whether the ability to perceive accurately a particular L2 contrast (e.g., the contrast between Spanish /p/ and /b/) is necessary for proper articulation of the contrasting phones. Flege's (1992) Speech Learning Model, for example, suggests that the inability to recognize or (re)categorize perceptual distinctions limits accurate L2 production. This paper will address that question with regard to the acquisition of the Spanish voiced and voiceless stop consonants by native English speakers.

The acquisition of the Spanish stops by English speakers is problematic for several reasons. While both languages contain a series of stop phonemes distinguished by voicing—voiceless /p t k/ contrasted with voiced /b d g/—the phonetic realization of the voiced and voiceless stops in the two languages differs in important respects. First, English /p t k/ are known as long-lag voiceless stops: they are realized with a relatively long voice onset time (VOT), which refers to the amount of time that elapses between the release burst of the stop and the onset of vocal fold vibration. This lag causes the aspiration that typically accompanies the production of /p t k/ in English, especially in word-initial position. The voiced stops, on the other hand, have short VOT values and are hence considered short-lag voiceless stops. This contrast is illustrated by the waveforms in Figures 1A and 1B, which show the first syllable of the words *poker* and *both*, produced by a native speaker of English, along with their respective VOT values.

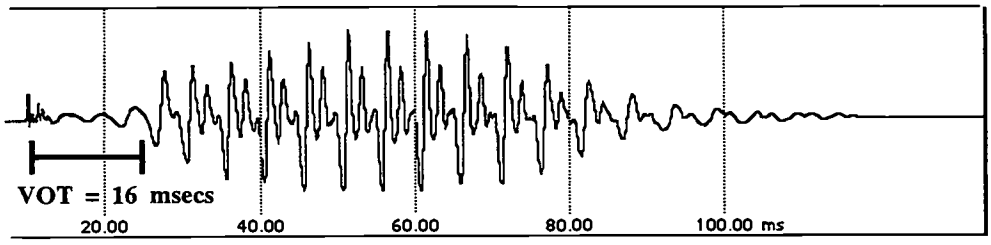
Unlike those in English, Spanish voiceless stops have short VOT values, while the voiced stops are realized with voicing lead (or prevoicing) in which



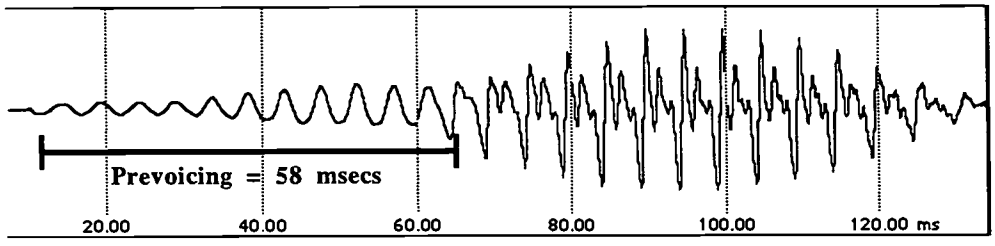
A. English *poker*



B. English *both*



C. Spanish *poca*



D. Spanish *boca*

Figure 1. Sample Waveforms for English /p/ and /b/ and Spanish /p/ and /b/

the vocal folds begin vibrating before the release burst. Waveforms illustrating this distinction for the Spanish words *poca* and *boca* (produced by a native speakers of Spanish) appear in Figures 1C and 1D. Thus, from a phonological perspective, /p/ and /b/ differ with regard to the feature [voiced] in both languages. From a phonetic perspective, however, Spanish /p/ is more like English /b/ in that both belong to the short-lag stop category and are characterized by short VOT values, as shown in Figures 1B and 1C of Figure 1. To illustrate the differences between the two languages further, Table 1 provides the average VOT production values for native speakers of Spanish and English, as well as the range of VOT values found, reported in a classic study by Lisker and Abramson (1964).

Given the phonetic differences between the two languages, the challenge for English-speaking learners of Spanish becomes clear: they must reorganize the phonetic categories of the voiced and voiceless stop phonemes so as to reflect those of the target language. The learner must shorten the relative VOT of /p/ during Spanish production, so that this phone falls within the target range of a short-lag stop, and also eliminate voicing lag from the production of Spanish /b/, so that this phone becomes prevoiced. In addition, the phonetic category overlap between Spanish /p/ and English /b/ has

Table 1
Mean VOT Measurements (in msecs) of English and Spanish Stops (Lisker & Abramson, 1964)

[Positive values indicate voicing lag; negative values indicate prevoicing]

English			Spanish		
Stop	Mean VOT	VOT Range	Stop	Mean VOT	VOT Range
/p/	58	20 / 120	/p/	4	0 / 15
/t/	0	30 / 105	/t/	9	0 / 15
/k/	80	50 / 135	/k/	29	15 / 55
/b/	1	0 / 5	/b/	-138	-235 / -60
/d/	5	0 / 25	/d/	-110	-170 / -75
/g/	21	0 / 35	/g/	-108	-165 / -45

important implications for Spanish L2 speech perception. Given their phonetic similarities, learners may confuse Spanish /p/ for /b/ perceptually and must therefore adjust their perceptual categories (or boundaries) to reflect the Spanish system in order to avoid confusion and promote comprehension.

METHODS

The differences in the phonetic realization of the Spanish and English stop consonants with regard to VOT provide the motivation for the current study. In particular, the study addresses the following research questions:

1. Do learners acquire the appropriate phonetic categories with regard to the Spanish voiceless and voiced stops? If so, how do they manipulate the different acoustic cues of the speech signal in order to achieve the necessary distinction?

2. What effect does formal training in phonetics have on the acquisition of the Spanish stops?

3. Is there a relationship between the perception and production of the Spanish stops? That is, do learners with native-like pronunciations of the Spanish stops also show evidence of native-like boundaries between the voiceless and voiced phonemes in perception?

While all three questions have some bearing on the results to be discussed, the present work focuses primarily on the issues raised by the third question, that of the relationship between production and perception in L2 Spanish acquisition.

In order to address the stated research questions, an experimental study was designed to examine both the production and perception of word-initial Spanish stop consonants by L2 learners whose native language was English. The participants were enrolled in an advanced undergraduate Spanish phonetics course at the University of Arizona. Thirteen learners volunteered for the study and were asked to complete a series of production and perception experiments in a speech laboratory setting at several points throughout the semester.

For the production portion of the experiment, each learner was recorded during four repetitions of an English or Spanish sentence, each containing a target word that began with a stop consonant. The English productions were obtained once during the second week of the semester, while Spanish productions were obtained three times. The first Spanish recording occurred during the third week of the semester, one week following the English production session. At that point, the learners had not yet begun to study the articulation of individual Spanish phonemes; instead, they had studied basic concepts of phonetics, as well as Spanish syllable structure and syllabification. The second Spanish recording session took place three weeks later, immediately after the Spanish voiceless stops had been studied in class. The text used for the presentation and practice of these phones was Barrutia and Schwegler (1994). In this text, the difference be-

tween the Spanish and English voiceless stops is not described in terms of short-lag vs. long-lag categories and VOT; however, the text does tell the learner to try to avoid aspiration of /p t k/ when speaking Spanish by maintaining greater muscular tension of the articulators and vocal tract during production. The final Spanish production experiment took place near the end of the semester--during the fifteenth week of class (and nine weeks after the second session). The sentences that the learners read were the same for all recording sessions. A total of 32 sentences were used with target words containing a variety of word-initial consonants. The data and results discussed in the present study come from the target stops that appear in sentences like those in Table 2.

Table 2
Sample Sentences Used in the Production Exercises

English:
Please say the word *paces* to me.
Please say the word *bases* to me.

Spanish:
Diga la palabra peso *por favor.*
'Say the word *peso* please.'

Diga la palabra beso *por favor.*
'Say the word *beso* please.'

Once all the production data had been gathered, the sentences were digitized, and the voiceless closure intervals and VOT of the target stops were measured from the digitized waveforms using SoundEdit 16. Voiceless closure interval refers to the duration of closure before the release of the stop consonant that is characterized by a lack of vocal cord vibration. Prevoicing of the voiced stops was also measured where applicable; in such cases, the duration of prevoicing was measured as a negative VOT (see Table 1).

Finally, spirantized variants of the voiced Spanish stops were noted as well, and these tokens were eliminated from the analysis. Spirantization is a process whereby Spanish /b d g/ are realized as approximants in certain phonetic contexts. The stop allophones generally appear after a nasal consonant and in phrase-initial position (and /d/ appears as a stop after laterals, as well), whereas the spirantized allophones appear elsewhere. Since the spirants do not have the closure, release burst, or VOT associated with stops, they were not included in the present analysis.

For the perception portion of the study, the learners listened to computer-edited versions of the English and Spanish nonsense words, *pada* and *bada*. These nonword tokens were natural speech tokens produced by a fluent English-Spanish bilingual, and they were edited so as to vary from 40 msecs of prevoicing to 56 msecs of voicing lag at approximately 5-msec intervals. The VOT continuum consisted of a total of 20 tokens for each language. The learners listened to both English and Spanish versions of the words presented randomly and indicated whether each one began with /p/ or /b/ by pressing the appropriate but-

ton on a response box in the testing room. Each token of the continuum appeared 10 times throughout the experiment for a total of 200 responses per language. This experiment was also carried out three times during the same three weeks as the Spanish production exercises: Weeks 3, 6, and 15 of the semester. In addition, 15 monolingual speakers of English and 12 Spanish-English bilinguals completed the perception experiment, so as to provide a basis of comparison for the L2 learners' perception data. A mean VOT perceptual boundary for /b/ and /p/ was determined for each subject group based upon the percentage of /b/ (or /p/) responses. Separate boundaries were obtained for the English and Spanish versions of the tokens.

RESULTS AND DISCUSSION

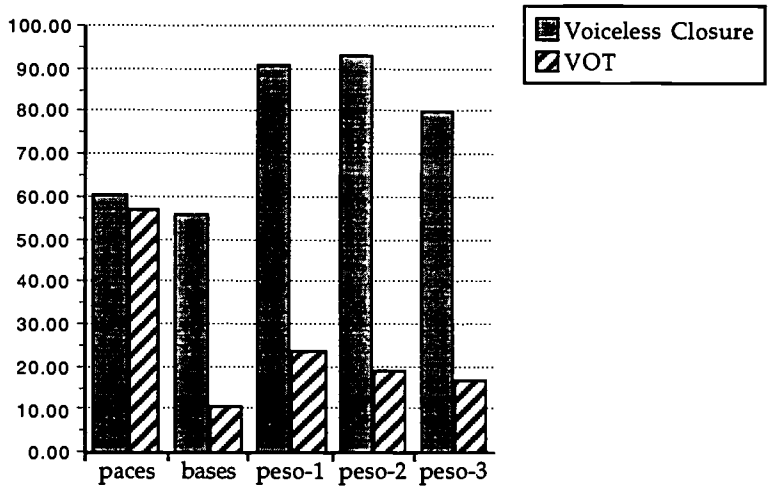
In the discussion of the results, the production results will be briefly described first, followed by a description of the perception results. Finally, the relationship between the learners' production and perception of Spanish /p/ and /b/ will be discussed.

Production Results

The production results are presented in Figures 2 and 3. These figures show the mean voiceless closure interval and VOT values for the English tokens *paces* and *bases* and the Spanish tokens *peso* ("weight") and *beso* ("kiss") for the L2 learner group as a whole.

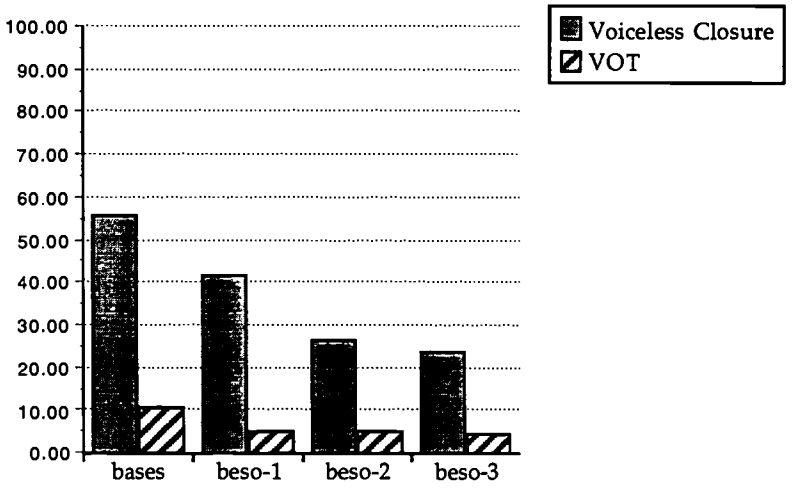
Consider first the information in Figure 2. As shown, the learners produced Spanish /p/ with VOT values that approach the average values reported for native Spanish speakers shown in Table 1. Although the learners' VOT values were somewhat longer than those of native Spanish speakers, they were still significantly shorter than their corresponding VOT values for English /p/. The learners also produced VOT values for Spanish /p/ that were similar to English /b/. Statistical testing of these data revealed that although the difference between the average VOT value of English /b/ in Figure 2 proved significant from that of *peso-1* and *peso-2*, there was no significant VOT difference between *bases* and *peso-3*. This indicates that by the end of the semester, the learners equated the L1 short-lag category, /b/, with the L2 short-lag category, /p/, at least with respect to VOT. Figure 2 also shows, however, that the learners distinguished Spanish /p/ from English /b/ through significantly longer closure intervals of the Spanish phone.

As for Spanish /b/, Figure 3 shows that the learners produced these tokens with somewhat shorter VOT values than in English; these differences, however, did not prove significant. Furthermore, the overall positive VOT means for L2 Spanish /b/ across all three sessions indicate that the learners failed to produce these tokens with prevoicing. As mentioned above, prevoicing is measured as a negative VOT; therefore, if the learners had prevoiced consistently (or learned to prevoice over the course of the semester), one would expect an overall negative VOT average. An examination of the individual data, however, revealed only two prevoiced /b/'s in the first session, three during the



peso-1: 1st session (Week 3) •*peso-2*: 2nd session (Week 6)*peso-3*: 3rd session (Week 15)

Figure 2. Mean Voiceless Closure and VOT Values (in msecs)—
Learner Production of English /p, b/ and Spanish /p/



beso-1: 1st session (Week 3) • *beso-2*: 2nd session (Week 6) • *beso-3*: 3rd session (Week 15)

Figure 3. Mean Voiceless Closure and VOT Values (in msecs)—
Learner Production of English and Spanish /b/

second session, and four during the third (out of a total of 56 tokens with word-initial /b/ in each session). Thus, the prevoicing associated with the Spanish voiced stops appears to take longer to acquire than the short-lag VOTs of the Spanish voiceless stops. (See Zampini (1998) for a detailed discussion of the production results presented in Figures 2 and 3, including a more detailed description of the statistical analyses.)

Perception Results

Turning now to the results of the perception experiments, consider first the information in Table 3, which shows the average VOT boundary that separates /p/ from /b/ perceptually for the L2 learners and the two control groups; Figure 4 presents the same information in graphical form.

First, it is interesting to note that all subject groups showed a consistent difference with regard to the VOT boundary for the Spanish and English versions of the nonsense words; namely, the Spanish boundary was consistently shorter than that for the English tokens. These two token types were identical in their acoustic characteristics prior to the release burst, so the voicing decision must

Table 3
Mean VOT Perceptual Boundaries (in msecs) Between /p/ and /b/

[Positive value indicates boundary in voicing lag range; negative value indicates boundary in prevoiced range.]

Group	English Tokens	Spanish Tokens	Mean
Monolingual English	14.5	2.77	**8.635**
Spanish/English Bilinguals	6.12	-8.96	**-1.42**
L2 Learners - 1st Session	8.22	-3.65	**2.285**
L2 Learners - 2nd Session	11.56	-7.59	**1.985**
L2 Learners - 3rd Session	11.85	-5.15	**3.35**

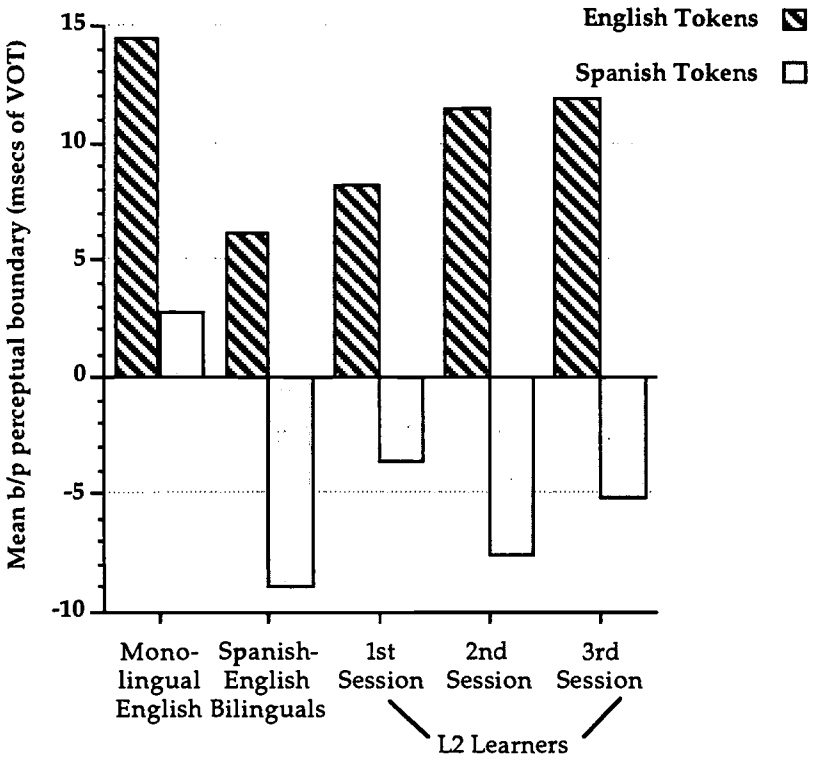


Figure 4. Mean VOT Perceptual Boundaries between /p/ and /b/

have been determined by aspects of the speech signal following the release burst; however, a discussion of the reasons for these differences is beyond the scope of the current paper.

Statistical analyses on the mean VOT boundaries revealed several significant effects. Consider first the boundaries for the English tokens in Figure 4. Analyses of these tokens revealed three significant differences. First, the English boundary difference between the monolingual English and Spanish-English bilingual speakers was significant. Second, there was also a significant difference in the English token boundaries for the monolingual English speakers and the L2 learners in the first L2 session, but not the second or third. Finally, the differences in English boundary between the L2 learners and the Spanish-English bilinguals were not significant for the first L2 session, but were for both the second and third sessions.

The Spanish token perceptual boundaries in Figure 4 also showed

several significant effects. First, the monolingual English speakers' Spanish boundary was significantly longer than that of the Spanish-English bilinguals; this pattern is similar to the one found for the English data of these two groups. Second, there was no significant difference in the Spanish perceptual boundaries of the monolingual English speakers compared to the L2 learners in the first session; however, the differences between these two groups in the second and third L2 sessions did prove significant. Third, none of the Spanish boundary differences between the L2 learners and Spanish-English bilinguals were significant. Finally, an examination of just the L2 learners' perceptual data revealed a significant difference in their English VOT boundaries between the first and second, and first and third, sessions, but not between the last two sessions. For the Spanish tokens, the only significant difference across sessions occurred between the first and second L2 sessions.

To summarize, the L2 learners started out with an English perceptual boundary that was significantly shorter than that of their monolingual English counterparts, but not significantly different from that of Spanish-English bilinguals; thus, their English boundary was more Spanish-like. These boundaries, however, shifted after training in the Spanish voiceless stops took place and became significantly longer and more English-like; this shift was sustained through the end of the semester. As for the Spanish perceptual boundaries, the L2 learners started out somewhere in between the two control groups, with a VOT boundary that was not significantly different from either group. They shifted toward more Spanish-like boundaries after training took place, however, as evidenced by the changes in the Spanish VOT boundary during the second session. Although this shift does not appear to be wholly sustained through the end of the semester, the L2 learners' Spanish boundary of the third session remained significantly different from the corresponding boundary of the monolingual English speakers, but not from the boundary of the Spanish-English bilinguals. Taken together, the changes in the L2 learners' English and Spanish perceptual boundaries indicate an attempt to maintain a clear distinction between the two languages by maximizing the perceptual distance between them. Thus, rather than showing evidence for one merged perceptual boundary between the voiced and voiceless stops regardless of language mode, the results suggest that the learners have two separate perceptual boundaries—one for each language—and that these boundaries become even more distinct with training.

The Relationship between Production and Perception

Finally, having discussed the individual production and perception results, we may now turn to the central issue of the paper: the relationship between the L2 learners' production and perception of the Spanish stops. If a positive relationship exists, one would expect that those learners who show short perceptual boundaries will also exhibit

short VOT production values, while those with longer perceptual boundaries will likewise exhibit longer VOT production values. Such an idealized relationship between production and perception is depicted in Figure 5.

In a similar fashion, if native-like perception in L2 is necessary for (or precedes) accurate production, one would expect that those learners with long perceptual boundaries will also have long VOT production values. Learners with short perceptual boundaries, however, may or may not have correspondingly short VOT production values. That is, if perception precedes production, learners that show accurate perceptual boundaries could still exhibit long production values if they are in a stage of acquisition in which production has not yet begun to change.

To examine this issue with regard to learner performance, correlations were obtained between the perceptual boundaries of a particular session and the corresponding productions of *peso* and *beso* from the same session. This information appears in Figures 6A - 6F.

In each of the graphs in Figure 6, the points represent the intersection of the Spanish perceptual boundary and mean Spanish VOT production value for each individual learner. None of the graphs approach the kind of idealized

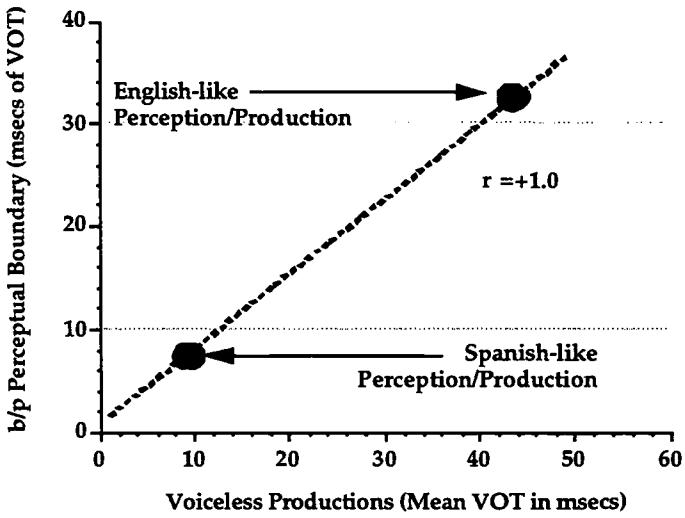


Figure 5. Idealized Relationship between Production and Perception

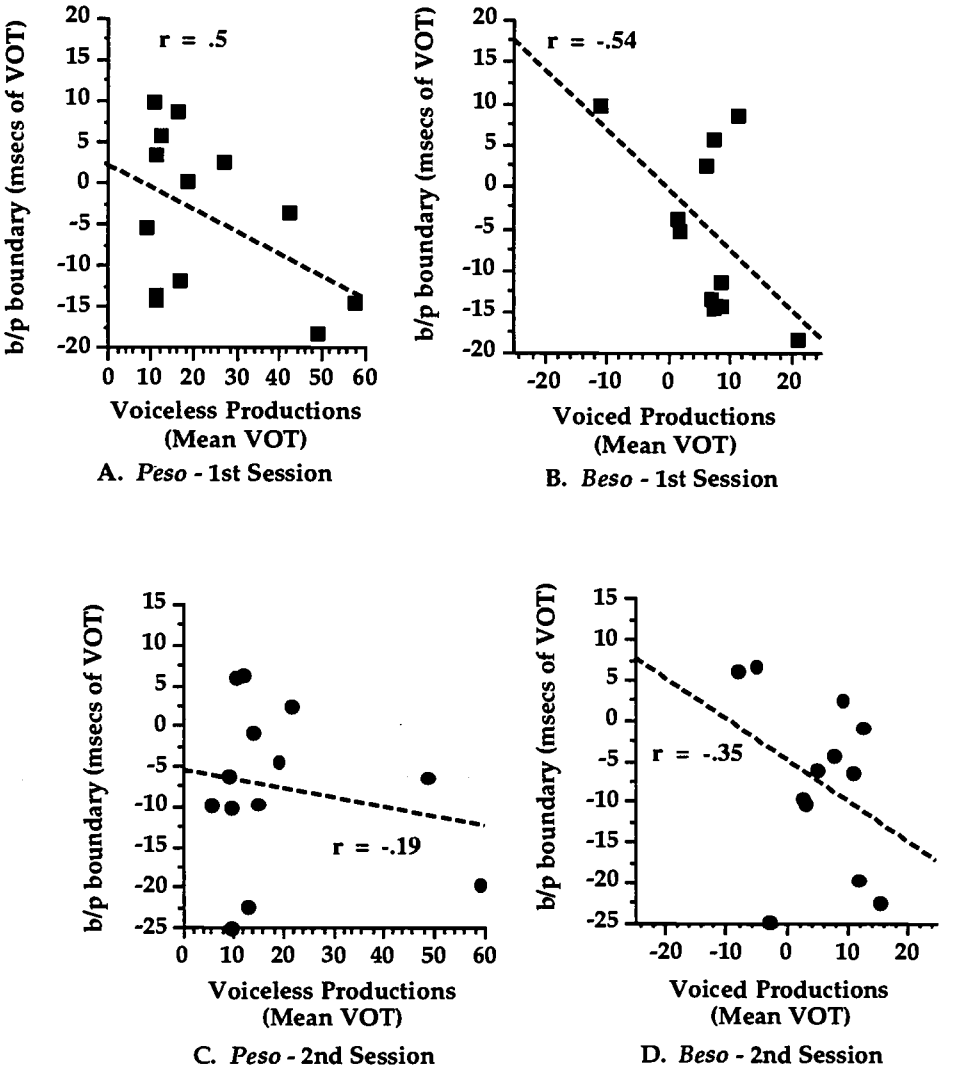


Figure 6. Correlations Between L2 Spanish Production and Perception (cont.)

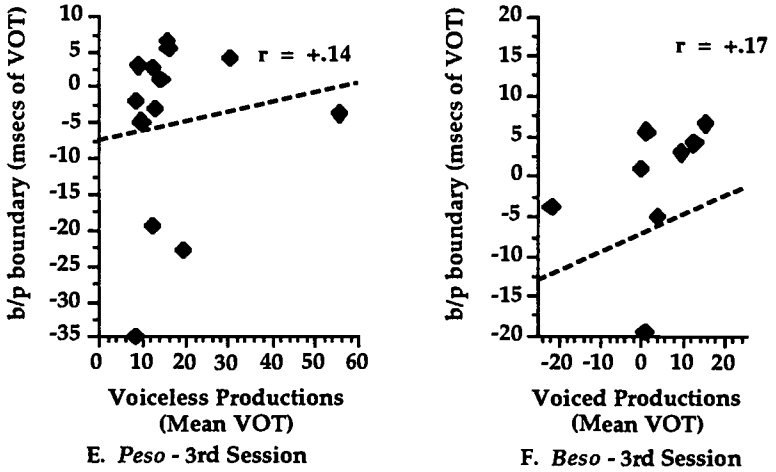


Figure 6. Correlations between L2 Spanish Production and Perception *(concluded)*

relationship shown in Figure 5. For example, consider the first session correlation data for the Spanish word *peso* in Figure 6A. Recall that if a positive relationship existed between production and perception, one would expect that learners with short VOT production values would also have short VOT perceptual boundaries. As seen in Figure 6A, however, those learners with the longest perceptual boundaries also have some of the shortest production values. This trend is also seen for the *peso* data in the second and third sessions, as shown in Figures 6C and 6E. These results, therefore, fail to support the hypothesis proposed by Flege's (1992) Speech Learning Model that inaccurate L2 perception will limit L2 production and suggest instead that perception does not necessarily precede production. In fact, the *peso* data in Figures 6A, 6C, and 6E appear to support an opposing hypothesis, namely, that L2 production may in some cases precede perception. That is, it may be the case that learners do not begin to adjust perceptual boundaries until they have attained accurate production categories. That this might be so is illustrated in Figure 7, which shows a predictive distribution of data for the opposing hypotheses.

If perception precedes production, one might expect a range of data along the horizontal axis, as depicted in Figure 7. Learners with short, Spanish-like perceptual boundaries may or may not exhibit correspondingly short production values, depending upon their stage of acquisition. If production precedes perception, on the other hand, one might expect a range of data along the vertical axis, since learners with short production values may or may not exhibit correspondingly short perceptual boundaries. This pattern

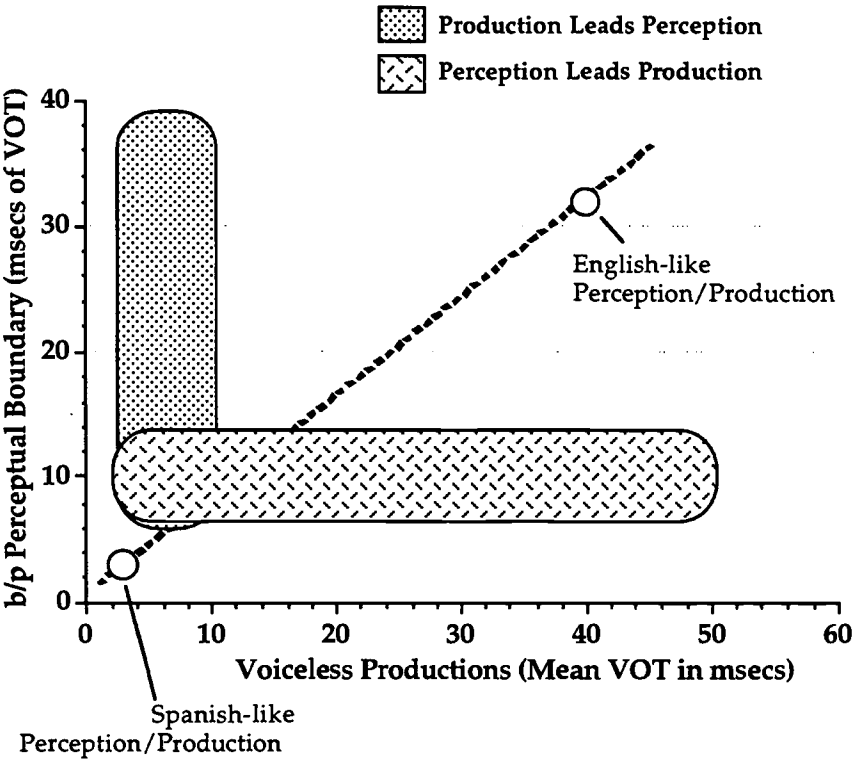


Figure 7. Relationship between Production and Perception - Two Hypotheses

would again depend upon each individual's stage of acquisition. Returning now to Figures 6A, C, and E, the overall spread of the individual points corresponds more closely to the predictions made by the hypothesis that production precedes perception. Thus, it appears that some learners learn to make the phonetic category substitution for Spanish /p/ before they make corresponding changes in perception. This result may not seem too surprising, since the phonetic category that they must learn for Spanish/p/, that of a short-lag stop, is one that already exists in the learners' first language. The phonetic category that learners must acquire for Spanish /b/, on the other hand, that of a prevoiced stop, does not exist in English. As a result, the substitution in production may take longer, which could, in turn, affect the interaction between production and perception. Consider, for example, the information in Figures 6B, 6D, and 6F, which correlate the Spanish perceptual boundary with the average VOT values for /b/. The correlations in these graphs are more scattered and do not appear to support either of

the predictions made by the hypotheses illustrated in Figure 7.

To summarize, the correlation data do not provide evidence for a positive relationship between the L2 production and perception of the Spanish stops. While some learners' L2 productions approximated those of native Spanish speakers, their perceptual boundaries were similar to boundaries for monolingual English speakers, and vice versa. In addition, the correlation data for Spanish /p/ suggest that production may precede perception, at least for this category. This indication does not necessarily mean, however, that inaccurate production will limit accurate perception; rather, it simply implies that the two may not be mutually dependent processes. The observed variation in the correlation data for Spanish /b/ further suggest that production and perception may be independent processes (at least for certain stages of acquisition or for certain types of phones), since no clear interaction was found.

These results have important implications for studies of both second language acquisition and speech processing. For example, although the results do not show a positive correlation between the perception and production with regard to VOT, there are other acoustic cues of the stop consonants that may play an influential role in the acquisition process. It was observed in Figure 2, for example, that learner production of the Spanish voiceless stops had significantly longer closure intervals than those for either the voiceless or voiced English stops; this difference provided a means of distinguishing the Spanish short-lag phones /p t k/ from the corresponding English ones, /b d g/. Given that the learners have learned to manipulate closure interval in order to achieve a given distinction, the interaction between VOT and closure may prove more important in L2 Spanish production and perception than either one of these two acoustic cues alone. Thus, research in both L2 production and perception is needed to gain additional insight into the relative importance of the different acoustic cues in second language acquisition, as well as into ways in which the weighting of these cues changes over time.

CONCLUSION

To conclude, this study has presented evidence for the acquisition of L2 Spanish stops by native English speakers as demonstrated by changes in production and perception over the course of the semester. Analyses that examined these processes individually showed significant changes toward Spanish-like production and perception categories over the course of the semester. Correlations of the production and perception results with regard to VOT, however, did not reveal a strong relationship between the two. The correlation data for Spanish /p/ provided some evidence for L2 acquisition in which accurate production precedes accurate perception, while the data for Spanish /b/ did not support either the notion that production precedes perception or that perception precedes production. It was suggested that the two processes may act independently during certain stages of acquisition. It may also be the case that, for some phones, per-

ception does precede production, while for others the reverse holds. Hence, this present study has served to illuminate a variety of potential interactions and to propose that the interaction between production and perception is more complex than is sometimes assumed.

REFERENCES

Barrutia, R., & Schwegler, A. (1994). *Fonética y fonología españolas*. New York: Wiley.

Flege, J. (1992). Speech learning in a second language. In C. A. Ferguson, L. Menn, & C. Stoel-Gammon (Eds.), *Phonological development: Models, research, and application* (pp. 565-604). Timonium, MD: York Press.

Lisker, L., & Abramson, A. (1964). A cross-language study of voicing in initial stops. *Word, 20*, 384-422.

Zampini, M. L. (1998). VOT and voiceless closure interval in the production of second language Spanish stop consonants. Manuscript submitted for publication.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

REPRODUCTION BASIS

☑ This document is covered by a signed "Reproduction Release (Blanket) form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

☐ This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").